

AD-A179 614

SUPERLINEAR CONVERGENT ALGORITHMS IN OPTIMAL CONTROL
(U) NORTH CAROLINA STATE UNIV AT RALEIGH E W SACHS
03 OCT 86 AFOSR-TR-87-0373 AFOSR-85-0214

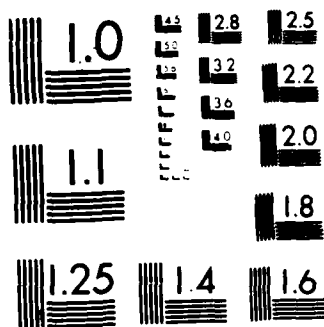
1/1

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DTIC FILE COPY

(2)

AD-A179 614

REPORT DOCUMENTATION PAGE

UNCLASSIFIED		1b. RESTRICTIVE MARKINGS									
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.									
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE											
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87 - 0373									
6a. NAME OF PERFORMING ORGANIZATION North Carolina State University	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research									
6c. ADDRESS (City, State and ZIP Code) Research Administration, Box 7033 Kaleigh, NC 27695-7003		7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448									
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFUSR-85-0214									
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448		10. SOURCE OF FUNDING NOS. <table border="1"> <tr> <th>PROGRAM ELEMENT NO.</th> <th>PROJECT NO.</th> <th>TASK NO.</th> <th>WORK UNIT NO.</th> </tr> <tr> <td>61102F</td> <td>2304</td> <td>15</td> <td></td> </tr> </table>		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.	61102F	2304	15	
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.								
61102F	2304	15									
11. TITLE (Include Security Classification) SUPERLINEAR CONVERGENT ALGORITHMS IN OPTIMAL CONTROL											
12. PERSONAL AUTHOR(S) Sachs, Ekkehard W.											
13a. TYPE OF REPORT Final Report	13b. TIME COVERED FROM 15/July/85 to 14/Sep/85	14. DATE OF REPORT (Yr., Mo., Day) 1986, 10, 3	15. PAGE COUNT								
16. SUPPLEMENTARY NOTATION											
17. COSATI CODES <table border="1"> <tr> <th>FIELD</th> <th>GROUP</th> <th>SUB. GR.</th> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table>		FIELD	GROUP	SUB. GR.				18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) A			
FIELD	GROUP	SUB. GR.									
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Problems involving the optimal control of ordinary or partial differential equations are infinite dimensional problems which are approximated by discretized problems for their numerical solution. Quasi Newton methods were applied to these finite dimensional problems and it was shown by analysis and numerical tests how the convergence rate could be predicted using information from the underlying infinite dimensional problem. For unconstrained optimal control problems with ordinary differential equations two approaches were studied: in one approach the control functions were used as unknowns whereas for the second route the control, state and costate functions were taken as unknowns. In the latter approach the quasi Newton update made extensive use of the special structure of the control problem and proved to be very effective. These algorithms were also studied for nonlinear elliptic boundary value problems and the optimal control of pseudoparabolic differential equations.											
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED									
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Sam Rankin		22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5020	22c. OFFICE SYMBOL NM								

APR 24 1987

2 AFOSR-TR-87-0373

1. Status of Research

1.1. Introduction

The research which was suggested in the proposal dealt with the application of quasi-Newton methods to optimal control problems. The main motivation consisted of the fact that these methods are very useful for optimization problems and exhibit a superlinear rate of convergence. This statement on the convergence rate was known to hold in infinite-dimensional spaces only under additional assumptions. Optimal control problems were formulated in infinite-dimensional spaces and hence the superlinear convergence behavior of quasi-Newton methods for these problems should be investigated.

1.2 Classical Quasi-Newton Methods

Optimal control problems of the following type were considered: Let $L: \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}$ and $f: \mathbb{R}^{n+m+1} \rightarrow \mathbb{R}^n$ for some $n, m \in \mathbb{R}$ and $T > 0, x_0 \in \mathbb{R}^n$.

Minimize

$$\int_0^T L(x(t), u(t), t) dt \quad (1)$$

subject to

$$\dot{x}(t) = f(x(t), u(t), t), \quad x(0) = x_0. \quad (2)$$

If it is assumed that (2) is uniquely solvable for all iterates u , then the objective can be written completely in terms of u

APR 24 1987

AIR FORCE SCIENTIFIC RESEARCH (AFSR)
NOTICE OF TECHNICAL RESEARCH
This technical report has been reviewed and is
approved for release to the public under AFM 130-12.
Distribution unlimited.
MATHEMATICS
Chief, Information Division

$$F(u) = \int_0^T L(x(u,t), u(t), t) dt. \quad (3)$$

The gradient of F is given by

$$\nabla F(u) = p^T(\cdot) f_u(x(\cdot), u(\cdot), \cdot) + L_u(x(\cdot), u(\cdot), \cdot), \quad (4)$$

where p solves the adjoint equation

$$\begin{aligned} -\dot{p}(t) &= p(t)^T f_x(x(t), u(t), t) + L_x(x(t), u(t), t), \\ p(T) &= 0. \end{aligned} \quad (5)$$

The computation of the Hessian of F is obviously even more complex so that the use of Quasi-Newton methods which do not require the knowledge of the Hessian is a desirable choice. Let $F: H \rightarrow \mathbb{R}$, H Hilbert space, be twice Fréchet-differentiable and $u_i \in H$ and $B_i \in L(H)$, the space of linear and bounded operators on H , be given. Then the BFGS method can be defined as follows:

- (i) Solve $B_i s_i = -\nabla F(u_i)$
- (ii) $u_{i+1} = u_i + s_i$
 $y_i = \nabla F(u_{i+1}) - \nabla F(u_i)$
- (iii) $B_{i+1} = B_i + \frac{\langle y_i, \cdot \rangle}{\langle y_i, s_i \rangle} y_i - \frac{\langle B_i s_i, \cdot \rangle}{\langle B_i s_i, s_i \rangle} B_i s_i.$

It was shown in [1] that for the control problem the superlinear rate of convergence in the Hilbert space norm



Distribution/	
Availability Codes	
Dist	Avail and/or Special
A1	

$$\lim_{i \rightarrow \infty} \frac{\|u_{i+1} - u_*\|}{\|u_i - u_*\|} = 0 \quad (7)$$

holds, if B_0 and u_0 are chosen close enough to $\nabla^2 F(u_*)$ and u_* and if, in addition

$$B_0 \approx H_{uu} + C, \quad (8)$$

where C is a compact operator and

$$H_{uu} v = (p_*(\cdot))^T f_{uu}(x_*(\cdot), u_*(\cdot), \cdot) + L_{uu}(x_*(\cdot), u_*(\cdot), \cdot)) v(\cdot).$$

Otherwise, one can expect at most a linear rate of convergence.

Obviously, the control problem (1), (2) cannot be solved numerically unless it is discretized. However, one might suspect, that the compactness condition (8) on the initial guess of the approximation of the Hessian influences also the convergence behavior of the discretized finite-dimensional problem. In [1] a fourth order Runge-Kutta scheme was applied with a Hermite interpolation at intermediate points. For the approximation of the inner product a composite Simpson's rule was used. Hence the discretized problem looked as follows:

For given $u^N \in \mathbb{R}^{2N+1}$ solve (1) and obtain $x^N \in \mathbb{R}^{2N+1}$.
 Then use (5) to compute $p^N \in \mathbb{R}^{2n+1}$ and evaluate (4) (9)
 at the grid points.

This procedure and the other possible route to discretize (1) and substitute it into (2) are quite different. In the latter case one obtains very complicated expressions for the gradient whereas the approach outlined in (9) is much easier to apply. However, (9) does not yield, in general, the gradient of a functional, so that the Jacobian is not symmetric and an application of the BFGS method seems not desirable because it maintains symmetry of the approximating matrices. But it was shown in [1] that this does not give rise to problems which is due to the fact that the finite dimensional problem is close to an infinite dimensional problem which exhibits all the features like self-adjointness and positive definiteness.

As a measure for the convergence in [1] the first iteration index was used for which the tolerance was achieved:

$$i(\epsilon) = \min\{i \in \mathbb{N} : \|\nabla F(u_i)\| < \epsilon\}$$

$$i_N(\epsilon) = \min\{i \in \mathbb{N} : \|\nabla F_N(u_i^N)\|_N < \epsilon\}.$$

Under appropriate conditions one can show, that for N large enough

$$i(\epsilon) - 1 \leq i_N(\epsilon) \leq i(\epsilon), \quad (10)$$

i.e. the termination criterion for the finite dimensional problems is asymptotically at the same iterate satisfied when it holds for the underlying infinite dimensional problem. Inequality (10) has been verified for numerical examples, but these tests were even more revealing with regard to the rate of convergence. In Fig. 1 a graph shows the decrease in the norm of the residual $\|\nabla F_N(u_i^N)\|_N$ for two choices of B_0 : For choice a) the compactness condition (8)

was satisfied, for choice b) it was not true. Both choices had approximately the same distance from the Jacobian. Although these two different initial selections for B_0 seem not to make a difference for the finite dimensional convergence behavior, Fig. 1 illustrates the fundamentally different convergence rate. This effect can be explained and even predicted by making use of the infinite dimensional theory.

1.3. Pointwise Approach

In [2] the same problem (1) and (2) was considered but the approach of applying quasi-Newton methods to it was quite different. Suppose, the necessary optimality conditions (2), (4) with $\nabla F(u) = 0$ and (5) are treated as a system of equations in the unknowns (p, x, u) . For $n = m = 1$ this means that one wants to solve

$$F(p, x, u) = \begin{pmatrix} \dot{x} - f(x, u, t) \\ \dot{p} + f_x(x, u, t)p + L_x(x, u, t) \\ f_u(x, u, t)p + L_u(x, u, t) \end{pmatrix} = 0$$

The Jacobian has the following structure in a proper function space:

$$F'(p, x, u) = \begin{pmatrix} 0 & D-f_x & -f_u \\ D+f_x & H_{xx} & H_{ux} \\ f_u & H_{xu} & H_{uu} \end{pmatrix}, \quad (11)$$

where $D = \frac{d}{dt}$ is the differential operator,

$$H(x,u,t) = f(x,u,t)p(t) + L(x,u,t),$$

and

$$F: X_r \rightarrow Y_r \text{ with } 1 \leq r \leq \infty$$

$$X_r = W^{1,r}[0,T] \times W^{1,r}[0,T] \times L^r[0,T],$$

$$Y_r = L^r[0,T] \times L^r[0,T] \times L^r[0,T].$$

Under a second order sufficiency condition for optimality one can show the regularity of F' and hence prove the quadratic rate of convergence for Newton's method. If a quasi-Newton method is designed for this problem one should take into account as much structure as possible. For example, the terms D , f_x and f_u in F' are known exactly because they are needed for the evaluation of F . Hence it is only necessary to update the lower right 2×2 block of F' in (11). All the operators in this block are multiplication operators so that an update should be performed with multiplication operators. This leads to updating the 2×2 -block for each t separately. In a discretized version this means that if u is replaced by a D -dimensional vector, then D 2×2 -blocks need to be updated. This can be a decisive advantage over the method presented in (1) where a $D \times D$ -matrix need to be updated, if D is large for the corresponding computing environment. Another perk of the pointwise updates is that at each step a linear two point boundary value problem needs to be solved, whereas in [1] the nonlinear state equation has to be solved at each iteration. The pointwise quasi-Newton method is described and analyzed in detail in [2] and has been also very successful for the numerical example. An interesting result is the rate of convergence which could be shown for the methods: with $z = (p,x,u)$ the following holds for all

$$1 \leq s < r \leq \infty$$

$$\lim_{i \rightarrow \infty} \frac{\|z_{i+1} - z_*\|_S}{\|z_i - z_*\|_r} = 0. \quad (12)$$

This is not the superlinear convergence rate (7) because different norms are used for $z_{i+1} - z_*$ and $z_i - z_*$, but these two norms can be arbitrarily close. This result (12) comes from the fact, that the system of nonlinear equations $F(p, x, u) = 0$ contains algebraic equations. This approach to solve optimal control problems shows a lot of potential for extension in direction of constraints and control of partial differential equations.

1.4 Pseudoparabolic Control Problems

In [3] a first approach towards optimal control problems with partial differential equations is undertaken. A model from heat conduction with memory was taken and a boundary control applied. The differential equation which is of pseudoparabolic type looks as follows:

$$\begin{aligned} y_t &= y_{xx} + \epsilon y_{xtx} & x \in (0,1), t \in (0,T) \\ y(0,x) &= 0 & x \in (0,1) \\ y_x(t,1) &= 0 & t \in (0,T] \\ -y_x(t,0) - \epsilon y_{xt}(t,0) &= u(t) & t \in (0,T]. \end{aligned}$$

Here $\epsilon > 0$ is a material constant and u is the control function. The solution of the differential equation can be represented by a Fourier expansion and is approximated by the first N terms of the series expansion. The theory developed in [1] can be applied to approximations of this problem and all the assumptions can be verified.

In the case that the objective functional is of an integral type,

$$\int_0^1 (y(T,x) - z(x))^p dx + \frac{a}{2} \int_0^T u(t)^2 dt$$

with $z \in C[0,1]$ and $a > 0$, $p \in \mathbb{N}$, then one can verify that the only nonlinearity is described by a real-valued function depending on real numbers. Based on the secant method one can construct an update which allows the same error estimates as the secant method and this yields a convergence rate of R-order $(\sqrt{5} + 1)/2$. Several numerical examples in [3] illustrate that this rate is actually observable. Similarly, a number of test runs were made for decreasing values of the constant a in the cost function. Then the problem became less and less well conditioned which resulted in a larger numbers of iterations to achieve the tolerance.

1.5 Elliptic Boundary Value Problems

Nonlinear elliptic boundary value problems are considered in [4]. Obviously, a discretization of this problem leads to a Jacobian with a large amount of sparsity. This is an advantage which should be used in the design of quasi-Newton methods. If one applies the Schubert algorithm to this problem, then from the tridiagonal structure of the Jacobian the update is given by [4, 1.12]. By letting the discretization parameters tend to the zero, one obtains as a continuous version.

$$(B_{i+1}v)(x) = (B_i v)(x) + (y_i - B_i s_i)(x)v(x)(s_i(x))^+, \quad (13)$$

$$\begin{aligned} y_i &= F(u_{i+1}) - F(u_i) \\ \alpha^+ &= \begin{cases} \frac{1}{\alpha} & \text{if } \alpha \neq 0 \\ 0 & \text{if } \alpha = 0 \end{cases} \\ F(u) &= \nabla^2 u + f(x, u, \nabla u). \end{aligned} \quad (14)$$

If one starts with an operator B_0 which includes the Laplacian, then the update (13) contains only multiplication operators. In the case where (14) does not depend on ∇u this approach works fine but otherwise the Jacobian is of the form

$$F'(u) = \nabla^2 + f_2(x, u, \nabla u) + f_3(x, u, \nabla u)\nabla,$$

which contains a derivative term of first order. This term is approximated by multiplication operators only and leads to problems in the numerical performance of Schubert's method. The remedy to this problem is to update B pointwise also for a derivative term:

$$\begin{aligned} (B_{i+1}v)(x) &= \\ & (B_i v)(x) + (y_i - B_i s_i)(x)(s(x)v(x) + \nabla s(x)\nabla v(x))(s(x)^2 + \nabla s(x)^2)^+. \end{aligned}$$

The fact that derivative terms are accounted for also shows in the convergence result:

$$\lim_{j \rightarrow \infty} \frac{\|u_{i+1} - u_*\|}{\|u_j - u_*\|} = 0 \quad \text{with } \|\cdot\| \text{ the } C^1\text{-norm}$$

The numerical results include nonlinear two-dimensional elliptic problems.

Also, the updates in this paper are pointwise quasi-Newton updates.

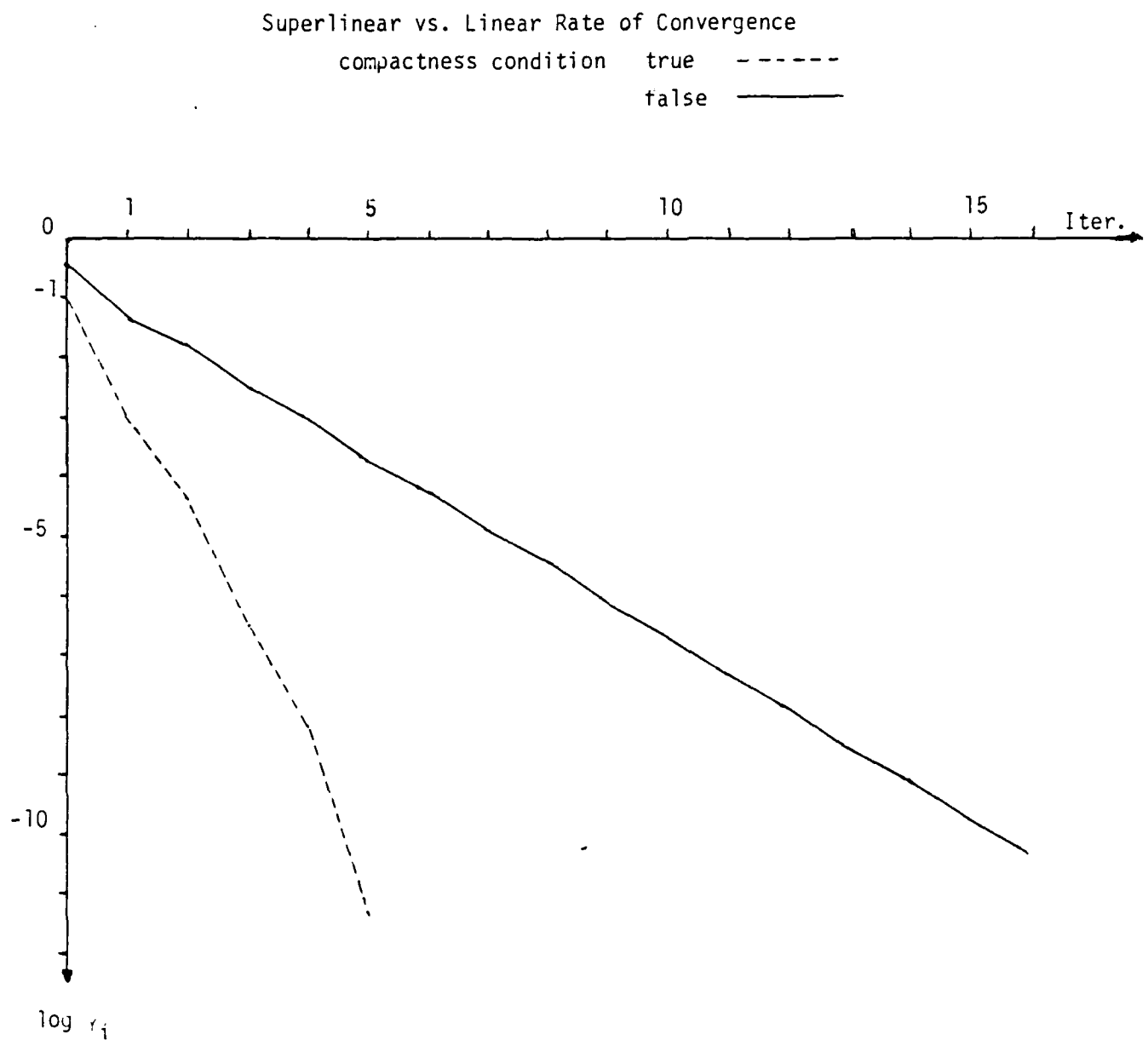


Figure 1.

Inventory Control Problem [1, p. 25 - 30]

$$\text{residual } r_i = \|g^H(u_i^N)\|_N, \quad N = 200$$

2. Publications

- [1] Quasi Newton Methods and Unconstrained Optimal Control Problems, (with C.T. Kelley), SIAM J. Contr. Optim. (submitted).
- [2] A Pointwise Quasi-Newton Method for Unconstrained Optimal Control Problems (with C.T. Kelley), submitted.
- [3] An Application of Quasi Newton Methods to Pseudoparabolic Optimal Control Problems (with C.T. Kelley), to appear in 'Proceedings on Optimal Control of Partial Differential Equations, Oberwolfach 1986', eds. K.H. Hoffmann and W. Krabs, Birkhauser, Boston.
- [4] A Quasi-Newton Method for Elliptic Boundary Value Problems (with C.T. Kelley), SIAM J. Num. Analysis (to appear).

3. Interactions

3.1 Invited Talks

Conference on Optimal Control of Partial Differential Equations,
Oberwolfach (West Germany), May 18-24, 1986.

Conference on Optimal Control and Calculus of Variations, Oberwolfach
(West Germany), June 15-20, 1986.

International Conference on Control and Identification of Distributed
Systems, Vorau (Austria), July 6-11, 1986.

Mathematics Colloquium, Pennsylvania State University, August 14, 1986.

3.2 Contributed Talks

12th International Symposium on Mathematical Programming, Boston, August
5-9, 1985.

SIAM Fall Meeting 1985, Tucson, October 28-30, 1985.

11th Symposium on Operations Research, Darmstadt (West Germany),
September 1-3, 1986.

3.3 Minisymposium

During the SIAM National Meeting 1986 in Boston, July 20-25, 1986, a
minisymposium was jointly organized by Tim Kelley and the P.I.. The topic was
'Quasi-Newton Methods in Infinite Dimensional Spaces' and the speakers were:

E. Allgower	(Colorado State University)
M. Trosset	(University of Arizona)
A. Griewank	(Southern Methodist University)
C.T. Kelley and E. Sachs	

END

5-87

DTIC